

Appl. No. : 09/539,032  
Filed : March 30, 2000

## AMENDMENTS TO THE CLAIMS

1. (Currently amended) A computer-based method for identifying conserved peptide motifs useful as drug targets for use in a host organism, wherein said method comprises the steps of:

- i) ~~generating computationally~~ computationally generating overlapping peptide sequences from selected organisms of length 'N',
- ii) ~~sorting computationally~~ computationally sorting the peptide sequences of length 'N' according to amino acid sequence,
- iii) ~~matching computationally~~ computationally matching the sorted peptide sequences of length 'N' of the selected organisms to produce matched common peptide sequences,
- iv) ~~locating computationally~~ computationally locating the matched common peptide sequences in the protein sequences of step i) and subsequently labeling the matched common peptide sequences with their origin and location;
- v) ~~joining computationally~~ computationally joining overlapping common peptide sequences to obtain extended conserved peptide sequences[[,]]; and
- vi) ~~annotating secondary structure of extended conserved peptide sequences based on a crystal structure database, and~~
- vii) ~~identifying conserved sequences not present in the host organism comparing said extended conserved peptide sequences obtained in step (v) to host organism protein sequences to determine which of said conserved peptide sequences are not present in host proteins, wherein said conserved peptide sequences which are not present in host proteins are useful as drug targets.~~

2. (Previously presented) The method of claim 1, wherein 'N' is at least 4.

3. (Previously presented) The method of claim 1 wherein the selected organisms include at least one of: Mycoplasma pneumoniae, Helicobacter pylori, Hemophilus influenzae, Mycobacterium tuberculosis, Mycoplasma genitalium, Bacillus subtilis, and Escherichia coli.

**Appl. No.** : **09/539,032**  
**Filed** : **March 30, 2000**

4. (Currently amended) A The method as claimed in of claim 1, wherein the conserved peptide motifs as modified comprising comprise sequences include one or more of the following sequences:

1. AAQSIGEPEGTQLT (SEQ ID NO:1)
2. AGDGTTTAT (SEQ ID NO:2)
3. AGRHGNKG (SEQ ID NO:3)
4. AHIDAGKTTT (SEQ ID NO:4)
5. CPIETPEG (SEQ ID NO:5)
6. DEPSIGLH (SEQ ID NO:6)
7. DEPTSALD (SEQ ID NO:7)
8. DEPTTALDVT (SEQ ID NO:8)
9. DHAGIATQ (SEQ ID NO:9)
10. DPHHGGGEG (SEQ ID NO:10)
11. DLGGGTFD (SEQ ID NO:11)
12. DVLDTWFS (SEQ ID NO:12)
13. ERERGITI (SEQ ID NO:13)
14. ERGITITSAAAT (SEQ ID NO:14)
15. ESRRIDNQLRGR (SEQ ID NO:15)
16. FSGGQRQR (SEQ ID NO:16)
17. GEPGVGKTA (SEQ ID NO:17)
18. GFDYLRDN (SEQ ID NO:18)
19. GHNLQEHS (SEQ ID NO:19)
20. GIDLGTTNS (SEQ ID NO:20)
21. GINLLREGLD (SEQ ID NO:21)
22. GIVGLPNVGKS (SEQ ID NO:22)
23. GKSSLLNA (SEQ ID NO:23)
24. GLTGRKIIIVDTYG (SEQ ID NO:24)
25. GPPGTGKTLA (SEQ ID NO:25)
26. GPPGVGKT (SEQ ID NO:26)
27. GSGKTTLL (SEQ ID NO:27)
28. GTRIFGPV (SEQ ID NO:28)
29. IDTPGHVDFT (SEQ ID NO:29)
30. ILAHIDHGKSTL (SEQ ID NO:30)
31. INGFGRIGR (SEQ ID NO:31)
32. IREGGRTVG (SEQ ID NO:32)
35. KMSKSKGN (SEQ ID NO:35)
36. KMSKSLGN (SEQ ID NO:36)
37. KNMITGAAQMDGAIL (SEQ ID NO:37)
38. KPNSALRK (SEQ ID NO:38)
39. LFGGAGVGKTV (SEQ ID NO:39)
40. LGPSGCGK (SEQ ID NO:40)
41. LHAGGKFD (SEQ ID NO:41)
42. LIDEARTPLIISG (SEQ ID NO:42)
43. LLNRAPTLH (SEQ ID NO:43)
44. LPDKAIDLIDE (SEQ ID NO:44)
45. LPGKLADC (SEQ ID NO:45)
46. LS GGQQQR (SEQ ID NO:46)
47. MGHVDHGKT (SEQ ID NO:47)
48. NADF DGDQMAVH (SEQ ID NO:48)
49. NGAGKSTL (SEQ ID NO:49)
50. NLLGKRVD (SEQ ID NO:50)
51. NTDAEGRL (SEQ ID NO:51)
52. PSAVGYQPTLA (SEQ ID NO:52)
53. QRVALARA (SEQ ID NO:53)
54. QRYKGLGEM (SEQ ID NO:54)
55. RDGLKPVHRR (SEQ ID NO:55)
56. SALDVSIQA (SEQ ID NO:56)
57. SGGLHGVG (SEQ ID NO:57)
58. SGSGKSSL (SEQ ID NO:58)
59. SGSGKSTL (SEQ ID NO:59)
60. SVFAGVGERTREGND (SEQ ID NO:60)
61. TGRTHQIRVH (SEQ ID NO:61)
62. TGVSGSGKS (SEQ ID NO:62)
63. TLSGGEAQRI (SEQ ID NO:63)
64. TNKYAEGYP (SEQ ID NO:64)
65. TPRSNPATY (SEQ ID NO:65)
66. VEGDSAGG (SEQ ID NO:66) and

**Appl. No.** : **09/539,032**  
**Filed** : **March 30, 2000**

33. IVGESGSGKS (SEQ ID NO:33)      67. VRKRPGMYIG (SEQ ID NO:67)  
34. KFSTYATWWI (SEQ ID NO:34)

5. (Currently amended) A The method as claimed in of claim 1, further comprising increasing the number of conserved peptide sequences by increasing the relatedness among the organisms being compared.

6. (Currently amended) A The method as claimed in of any one of claims 1-4 wherein the invariant conserved peptide sequences belong to are found within the sequences of at least one of the following proteins:

- I DNA DIRECTED RNA POLYMERASE BETA CHAIN
- II EXONUCLEASE ABC SUBUNIT A
- III EXONUCLEASE ABC SUBUNIT B
- IV DNA GYRASE SUBUNIT B
- V ATP SYNTHASE BETA CHAIN
- VI S-ADENOSYLMETHIONINE SYNTHETASE
- VII GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE
- VIII ELONGATION FACTOR G (EF-G)
- IX ELONGATION FACTOR TU (EF-TU)
- X 30S RIBOSOMAL PROTEIN S12
- XI 50S RIBOSOMAL PROTEIN L12
- XII 50S RIBOSOMAL PROTEIN L14
- XIII VALYL tRNA SYNTHETASE
- XIV CELL DIVISION PROTEIN FtSH HOMOLOG
- XV DnaK PROTEIN (HSP70)
- XVI GTP BINDING PROTEIN LepA
- XVII TRANSPORTER and
- XVIII OLIGOPEPTIDE TRANSPORT ATP BINDING PROTEIN OPPF.

Appl. No. : 09/539,032  
Filed : March 30, 2000

7. (Currently amended) A The method as claimed in of claim 1, wherein ~~the said method of comparing the peptide libraries as given in step (iii) of claim 1 is carried out by following the steps~~ comprises:

selecting organism names from a menu;

iteratively comparing peptide sequences of a first organism to peptide sequences of a second organism; and, for matching sequences, writing sequences to a first file for the first organism and to a second file for the second organism.

8. (Currently amended) A The method as claimed in of claim 1 wherein ~~the said method of locating the common peptides in the original protein sequences as given in step (iv) of claim 1 is carried out by the following steps~~ comprises:

selecting protein sequences;

iteratively comparing matched peptide sequences to protein sequences; and

if the peptide is found in a protein sequence, labeling the peptide sequence in a file associated with the protein with: a) a protein identification number (PID), b) a location in the protein sequence, and c) a name of the organism.

9. (Currently amended) A The method as claimed in of claim 1, wherein ~~the said method of creating a common peptide of variable length after removing the overlapping as given in step (v) of claim 1 is carried out by the following steps~~ comprises:

iteratively comparing data on matched peptide locations;

determining overlapping matched peptides; and

determining extended peptide sequences based on overlapping matched peptide sequences.

10-12. (Canceled)